

实证经济学近三十年进展之我见

中国人民大学汉青经济与金融高级研究院 刘延洁

北京大学国家发展研究院 赵洪春*

上世纪八十年代初以来的三十年间,实证经济学在可信度上取得了很多意义深远的进展。除了得益于更多更好的数据和更强大的计算能力之外,经济学家也有了更精良的分析工具,以及对相关理论问题更深刻的理解。这些变化集中体现在基于设计的实验主义实证研究方法、稳健的经济计量推断和基于经济学模型的结构性质实证研究之中。其中,实验主义方法是实证经济学过去三十年来取得的最引人瞩目的进展。这些进步都提高了实证研究的可信度,因此可以说:三十年来,实证经济学最重要的进展是可信度的提高。

过去,造成实证研究不可信的主要原因有二:一是因果关系无法识别,二是统计推断不稳健。实验主义方法,连同稳健的推断方法,令人信服地解决了在特定情形下因果效应有无的问题。虽然如此,因果关系背后的机制仍不能有效识别,并且在处理一般均衡问题和涉及策略行为的问题时需要结构性实证方法。本文即试图梳理归纳这些主要进展。

1 基于设计的实验主义实证研究方法

实验主义实证研究方法可以追溯到 Lalonde (1986) 的研究。当时盛行的政策评估方法多是基于计量经济学的;但在一些评估美国政府补贴的培训项目是否有用的研究中也应用了随机试验。Lalonde (1986) 发现上述两种方法对同一项目的评估结果相差甚远,进而质疑统计性分析能否为政策评估研究指明正确方向。在总结了一系列相关研究之后, Ashenfelter (1987) 认为随机试验方法是政策评估研究应采用的路线。之后,使用自然实验讨论劳动经济学和公共财务问题的研究大量涌现,并且直接推动了实验主义方法的发展。Griliches (1986) 在《计量经济学手册》一书中曾经说:“如果数据是完美的,即是从设计良好的随机试验中获得的,那么就没有计量经济学这个学科了。”实验主义方法清晰简洁,便于直接锁定关于各自设定是否正当的具体问题,有自己独特的优势。目前,已经在诸如发展、教育、环境经济学、卫生、劳动和公

* 刘延洁,现为中国人民大学汉青经济与金融高级研究院研究生,电子邮箱:liuyanjie@ruc.edu.cn。赵洪春,获得美国南加州大学经济学博士学位,现为北京大学国家发展研究院助理研究员,电子邮箱:zhaohongchun@nsd.pku.edu.cn。

共财务等领域得到广泛应用。^①

最严格的实验主义方法是随机控制试验(Randomized Control Trials, RCTs)。这种方法随机地将样本分为实验组和对照组,并把某种待研究的措施或者政策仅施加在实验组样本上,然后计算出两组中被解释变量的平均差异,即某措施在某个特定情境下对被解释变量的平均干预效果(Average Treatment Effect, ATE)。这样做的好处是直接克服了回归分析中常见的内生性问题,且不需要复杂的计量经济学技巧即可直接估计平均干预效果。在做随机控制试验时需要特别注意的是:分派实验措施时是否真正随机,以及实验前的变量在实验组和对照组之间是否平衡。

一个早期的随机试验是1974年的兰德健康保险实验(RAND Health Insurance Experiment)(实验的详情参见Manning, Newhouse, Duan, Keeler 和 Leibowitz, 1987)。但是很多早期的随机试验都有严重缺陷,比如,负收入税实验(如Seattle/Denver and Gary Income Maintenance Experiments)就存在严重的样本丢失(sample attrition)和系统性地误报收入情况。现在,由于有了行政数据,以及在没有行政数据时,能更精确地解读调查数据,这些问题已经可以解决。迄今,随机试验方法已应用于多个微观经济学领域。比如,在政策评估研究中有墨西哥的Progresa项目(详情参见Gertler, 2004; Schultz, 2004),美国的Moving to Opportunity项目(详情参见Kling, Liebman 和 Katz, 2007);在估计结构性参数的研究中,Fehr 和 Goette(2007)使用该方法估计跨期替代弹性;在发展经济学研究中,有Banerjee, Duflo, Glennerster 和 Kinnan(2009)对印度小额贷款的研究,和Miguel 和 Kremer(2004)对肯尼亚肠道寄生虫的研究。

随机控制实验费钱费时,且有时候并不可行。当无法实现随机控制试验时,可以采用自然实验来识别解释变量与被解释变量间的因果效应。自然实验包括工具变量法(Instrument Variables, IV)、二阶差分法(Differences in Differences, DID)和断点回归法(Regression Discontinuity Design, RDD),这些自然实验方法也是严谨可取的实验主义方法。

工具变量是那些在系统中只影响解释变量的外生变量。借助源于工具变量的外生冲击,可以识别解释变量对被解释变量的因果效应。使用这种方法时需要特别注意除了通过内生变

① 目前,实验主义方法在宏观经济学和产业组织中的应用前景仍不明朗,不过,其在这些领域的应用已经有一些有益的探索。比如货币政策冲击对实体经济的影响(Romer 和 Romer, 1989, 2004; Richardson 和 Troost, 2009; Velde, 2009),财政政策冲击对实体经济的影响(Ramey 和 Shapiro, 1998; Johnson, Parker 和 Souleles, 2006; Romer 和 Romer, 2010; Ramey, 2011; Barro 和 Redlick, 2011),经济增长的影响因素(Acemoglu, Johnson 和 Robinson, 2001; Bleakley, 2007),并购对价格的影响(Hastings, 2004; Ashenfelter 和 Hosken, 2010)和金融市场的扭曲(Keys, Mukherjee, Seru 和 Vig, 2010; Bubb 和 Kaufman, 2009)。

量这个渠道之外,工具变量本身对被解释变量应没有任何影响。^①

二阶差分法比较受某个事件影响的“实验组”和不受该事件影响的“对照组”在事件发生前后的平均变化,并将其差异解释为因果效应。使用这种方法时需要特别注意的是两组必须有相同的变动趋势;如果有各自特有的变动趋势,那么这种方法就无法识别因果效应。二阶差分法在经济学中应用广泛,比如,Card(1990)研究外国移民对失业率的影响;Hastings(2004)研究并购对价格的影响;Rodrik 和 Wacziarg(2005),Persson 和 Tabellini(2008)研究民主对经济增长的影响。

当某个临界点能够将人群自然分组,在临界点上方和下方的不同组会突然地受到不同的干预时,可以使用断点回归法。这种方法可以视为一种特殊的工具变量法,不过此处的工具变量是由控制变量和内生变量间的非线性或者不连续关系直接推出的。使用这种方法的有影响的实证研究有 Angrist 和 Lavy(1999)对班级大小和学习成绩关系的研究;Hahn, Todd 和 van der Klaauw(2001)对反歧视法的研究;van der Klaauw(2002)对奖学金的研究以及 Lee(2008)对美国众议院选举的研究。使用这种方法时需要特别注意的是断点附近是否聚集了大量的样本;如果聚集了大量样本,那么表示断点直接影响了当事人的行为,从而无法估计因果效应。

实验主义方法不仅改变了众多微观领域中的实证研究形式,而且在处理一些受到广泛关注的实际问题时也取得了巨大的成功。其中一个著名案例是美国 1970 年代在是否执行死刑上的反复,以及由此引发的关于死刑能否吓阻谋杀的讨论。美国在 1972 年到 1976 年间事实上停止执行死刑,但是在 1977 年 1 月恢复了死刑。之所以恢复死刑,Isaac Ehrlich(1975a, 1977b)的一系列研究起到了推动作用。在这些文章中,作者使用工具变量法发现死刑对谋杀具有切实的吓阻作用。不过,这些文章也受到猛烈的批评。比如,Leamer(1983)使用敏感性分析有力地说明这种吓阻作用至多是弱的。就此问题,双方争执不下,一时难有定论(参见 Bowers 和 Pierce, 1975; Passel 和 Taylor, 1977; Ehrlich, 1975b, 1977a; Ehrlich 和 Liu, 1999)。这一类研究的要害在于死刑和谋杀案件的发生是相互决定的,而且其他遗漏变量也会改变从死刑到谋杀的因果关系。

较可信地解决这个问题的是 Donohue 和 Wolfers(2005)的研究。作者使用二阶差分法对比了美国和加拿大年度谋杀率的时间序列数据。研究表明尽管在美国死刑政策发生了迅速的

① 采用工具变量法的实证研究浩如烟海,部分在劳动、卫生经济学中的代表性工作有 Angrist 和 Krueger(1991), Card(1995), Duflo(2001)对教育回报的研究; Angrist(1990), Imbens 和 van der Klaauw(1995)对退伍军人收入的研究; Anderson 和 Meyer(2000), Gruber(2000)对社会保障的研究; Bronars 和 Grogger(1994), Angrist 和 Evans(1998)对生育和劳动供给关系的研究; McClellan, McNeil 和 Newhouse(1994)对心脏病手术条件和健康关系的研究; Evans 和 Ringel(1999)对母亲是否吸烟和婴儿出生时体重关系的研究; Levitt(1997)对警力和犯罪关系的研究。

变化,但是这个变化对谋杀率的影响微乎其微。同时,由于年度谋杀率波动较大,吓阻效应即便存在也必须足够大才有可能被识别到。

实验主义方法解决的另一个重要实际问题是估计教育的生产函数。这一类文献始于 Coleman 等人(1966)的报告。该报告应用回归中的拟合优度分析各种教育投入能在多大程度上解释学习成绩。Coleman 报告和许多随后的回归分析都发现教育投入和学习成绩之间的关系脆弱。此类研究的主要问题仍然是反向因果关系(reverse causality)和遗漏变量偏差(omitted variables bias)。具体而言,是能否把教育投入与其他复杂因素——比如,学生、学校和社区的特征——区分开来。

Card 和 Krueger(1992a,b)是较早使用工具变量法解决这个问题的研究。作者隐含地将美国各州在不同年代的教育支出变化视为一个自然实验,采用各州各个出生年代虚拟变量作为工具变量,估计了教育投入对教育回报的影响。研究表明那些出生在教育支出更高的州和年代的个人有更高的教育回报。

Angrist 和 Lavy(1999)依据以色列的 Maimonides 法则构造了基于断点回归法的工具变量来研究班级大小对学生成绩的影响。结果证实班级规模减小显著提高了成绩。美国田纳西州 STAR 实验是应用随机控制试验研究教育生产问题的典型案例。Krueger(1999)根据这项实验发现即便控制了样本丢失和自选择问题,小班仍然对学习成绩有显著的益处,进一步证实了此前的发现。

2 稳健的经济计量推断

计量经济学中,为了通过数据识别因果效应往往需要技术性的附设假设,而稳健的经济计量推断尽可能地降低了各种附设假设对计量分析结果的影响。

例如,在截面回归分析中计算标准误通常假设误差项是同方差的。这里的同方差假设是一个附设假设,因为即使是异方差,估计值的性质也不受影响,只是置信区间不再可靠罢了。1970 年代处理异方差的常用方法是先检验是否存在异方差;如果存在,则将方差视为解释变量的函数,然后使用加权最小二乘法(Weighted Least Squares)估计系数(参见 Gujarati,1978)。但是,这种方法使用不便,且会影响到参数的估计值,因而应用不广。White(1980)提出了如何计算不论是否存在异方差时都有效的标准误。这种稳健的标准误(robust standard errors)已成为计量经济学的标准工具。此外,在面板数据分析中,由于误差项可能是序列相关的,所以传统的稳健标准误也不可靠(Bertrand, Duflo 和 Mullainathan,2004)。可以证明,面板数据中,如果误差项在某些单元间不相关,那么丛聚的标准误(clustered standard errors)对单元内部的异

方差和序列相关是稳健的。通常,截面单位数量远超时序单位数量,比如家庭数远远大于时点数。在这种情形下,丛聚的标准误仍是一致的。最近 Hansen (2007) 发现,即便时序单位数量远超截面单位数量,有效推断仍然可行。

稳健推断的另一个例子是弱工具变量下的推断。为保证实证结果的可信度,采用工具变量法时优先关注的是保证工具变量与误差项不相关。这样做的一个副作用是:在控制了其他变量后,工具变量往往只与其处理的内生变量弱相关。弱工具变量使得推断变得困难,即便是在大样本情况下,弱工具变量也使得检验统计量的渐进分布不能近似其抽样分布。理论计量经济学家已经发展出一整套弱工具变量情况下的统计推断工具。其中最简单的是看第一阶段回归的 F-统计量,即在两阶段最小二乘法的第一阶段检验所有工具变量的影响为零这一假设。如果 F-统计量大于 10,那么就可以认为不存在弱工具变量,并且可以使用常规的两阶段最小二乘法;否则不可以。弱工具变量的替代方法众多,比如,在只有一个内生变量时,不论工具变量是强是弱还是错,Moreira (2003) 提出的条件似然率统计量都可以构造有效且最优的置信区间和假设检验。此外,有限信息最大似然估计方法 (Limited Information Maximum Likelihood, LIML) 在很多情况下的小样本表现要优于弱工具变量下的两阶段最小二乘法。

实验主义研究方法也激发了对主要解释变量与控制变量在回归中具有不同地位的讨论。传统的教科书中假设所有解释变量,不论是主要解释变量还是控制变量,都和误差项无关。在这种情况下,所有解释变量的回归估计系数都是无偏的,都可以解释为因果关系。这个假设在很多情况下过强了。在实验主义方法的框架下,主要解释变量和控制变量之间的区别是清楚的。相应的识别假设为较弱的条件期望独立 (Conditional Mean Independence)。举例来说,在考虑班级大小与学生成绩的关系时,除了班级大小外,很多其他难以直接度量的因素也会影响成绩,比如父母辅导孩子的时间。在回归中,可以把所在地区的收入纳入回归,作为这些遗漏变量的代理,因此地区收入自然与误差项相关。这表示此时“误差项独立”的假设不合理。其实,只要班级大小与控制了地区收入后的误差项无关,即满足识别条件。在这个较松的假设下,班级大小的估计值可作因果关系解释,尽管地区收入的估计值不可以。

最后提一下非参数和半参数方法 (Powell, 1994)。目前多数回归分析采用线性函数形式,能否在放松函数形式后进行推断呢? 非参数和半参数估计就是回答这个问题的,尤其是半参数推断更重要。该方法在很弱的附设函数形式假设下即能够可信地识别因果关系。非参数回归在直观上容易理解,在不做任何函数形式的设定下,选取在解释变量特定取值附近的样本计算相应区间被解释变量的平均值,作为解释变量取该值时被解释变量的条件期望。非参数和半参数估计都需要大量的数据才可行,这多少限制了这两种方法的应用。

3 基于经济学理论的结构性实证研究

尽管实验主义方法和稳健的统计推断已经成为实证经济学家的有力工具,但是它们与经济学理论的联系仍是间接的。如何精确地匹配经济学理论与观察到的数据是近三十年来实证经济学的另一个发展方向。尤其是在产业组织和宏观经济学中,明确使用经济学模型进行数据分析的结构性实证研究日益成为主流。这些研究尝试使用数据来识别相关经济学模型中的参数,比如用来刻画偏好和技术的参数。与实验主义方法相比,结构性实证研究种类多样,涵盖范围和使用技巧颇广,在不同领域中与经济学理论的联系程度也颇不相同。比如,从使用的经济学理论看,具体模型可以是基于个人的最优选择,也可以是基于从个人选择导出的加总关系;从和经济学理论结合的程度看,有时采用参数化的定量模型,有时仅采用一般的定性结论;在估计结构性参数时,又往往揉合了校准方法和回归方法。总的来说,与实验主义方法相比,结构性实证研究的优势不是识别因果效应是否存在,而是预测和政策评估。因此,二者在很大程度上是互补的。下面举几个例子来解释它的一些特点。

产业组织中的一个重要问题是研究并购对价格的影响及其背后的具体机制。研究并购的主要结构性研究范式被称为新实证产业组织(new empirical industrial organization),它包括三个步骤:

第一步是估计所研究产品的需求函数系统,通常采用 Berry, Levinsohn 和 Pakes (1995) 发展的离散选择和差异产品框架。在估计需求弹性时使用工具变量法,通常选择其他市场的价格作为工具变量。

第二步是构造一个市场结构模型来刻画市场上各厂商的竞争关系,比如含有不同品牌和产品的 Bertrand 价格竞争模型。在此模型中,由前一步算出的替代弹性和厂商的利润最大化行为可以推出一系列价格和边际成本之间的关系。

最后,模拟该产业在发生并购和没有发生并购情况下的行为。可以看出,新实证产业组织首先将理论和观察到的行为匹配,并且从中揭示有关偏好和技术的参数;然后用这个模型估计对各种情况的反应,包括从未发生的情形。Nevo (2000) 即严格采用了该范式来研究并购对早餐麦片价格的影响。^①

结构性实证研究的另一个重要的应用领域是宏观经济学。宏观实证经济学考虑的问题大致可以分为三类:(1)估计结构模型中的参数;(2)当政策规则、制度或者偏好发生改变时,宏

① 类似地,在市场营销文献中 Erdem, Imai 和 Keane (2003) 使用基于模型的方法将降价促销效应分解为三个具体的机制:品牌扩张的部分、从其他品牌转移来的部分和提前消费的部分。

观经济的动态结构如何变化;(3)在当下的制度和政策规则中,一次独立的外生冲击或政策冲击对内生变量的影响。前两类问题都必须用到结构性实证研究。

正如 Sims(1980)所指出的,由于受到多方面的限制,早期的宏观实证研究存在可信性问题。同时又由于只有一部分问题可以用实验主义方法,现代宏观经济学发展出了与微观经济学不同的方法来研究实证问题。比如,向量自回归(vector autoregressions, VAR),结构向量自回归(structural vector autoregressions, SVAR)和动态随机一般均衡(dynamic stochastic general equilibrium, DSGE)模型。可以看出,计量方法和经济学理论的有机结合日益重要。

举例来说,宏观经济学中的一个重要问题是货币政策对实体经济的影响。这类研究的难点在于将货币政策与中央银行对非政策扰动的回应相分离。比如, Romer 和 Romer(1989)在其开创性地将实验主义方法引入货币政策分析的研究中,就未能分离二者。结构向量自回归模型在一定程度上解决了这个问题。这种技术把根据理论推导得出的少许约束与多变量时间序列模型相结合,可以将政策效应表达为由回归估计值的函数。因为只使用少量定性推论,所以其数据拟合程度与未作任何约束时几无差别。结构向量自回归模型中所做假设之少令人惊奇。常见的假设有利率变化对产出和通胀的影响有一个月或者一个季度的滞后期,或紧缩的货币政策至少不会推高产出和价格水平,又或货币政策在长期不改变真实变量。Sims 和 Zha(2006)即在此框架下重新审视了战后的货币政策,发现货币政策变量的变动代表的是政策对经济状态的反应,而不是外生冲击。经济的周期变化几乎都不能归因于货币政策的变化。此外,他们的结论显示货币政策的实际效应比通常认为的要小很多,不确定性也更大。

尽管兴起不久,动态随机一般均衡模型已经成为货币政策研究的利器(Smets 和 Wouters, 2003)。与结构向量自回归模型比,动态随机一般均衡模型有更强的假设,因而也有更强大的解释力,而且在拟合数据方面也不逊色。该模型的优势在于,作为一个讨论政策的框架,它可以明白揭示政策影响经济的机制,也能解释非政策冲击的原委。在对待参数的不确定性方面,最先进的动态随机一般均衡模型采用贝叶斯(Bayesian)方法来推断参数的取值,并且在很大程度上取得了成功。

4 小结

实证经济学在过去三十年间取得了巨大进展。面对上世纪八十年代的诸多挑战,实证经济学家发展和改进了实验主义方法、稳健的经济计量推断和结构性实证研究,提高了可信地推断因果关系的能力。这些新发展不仅改变了计量经济学理论和众多领域中实证研究的面貌,而且加深了对理论和数据之间关系的理解,也厘清了实证经济学中已经解决和很多亟待解决

的问题。

虽然实证经济学发展迅速,但是在可信地推断经济行为背后的因果关系方面仍未克竟全功。在此方面,不妨将计量方法与定量方法结合起来。具体而言,是用参数化的定量模型构造变量之间的因果关系,同时使用包括实验主义方法在内的归纳方法来估计模型中的结构性参数,以及构造相关的却无法直接观察的变量。可以预期,未来实证经济学中将有更多采用各种参数化定量模型的结构性实证研究。

参考文献:

- Acemoglu, Daron, Simon Johnson, and James A. Robinson, 2001, “The Colonial Origins of Comparative Development: An Empirical Investigation”, *American Economic Review*, 91(5), 1369–1401.
- Anderson, Patricia, and Bruce Meyer, 2000, “The Effects of the Unemployment Insurance Payroll Tax on Wages, Employment, Claims and Denials”, *Journal of Public Economics*, 78(1–2), 81–106.
- Angrist, Joshua D., 1990, “Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records”, *American Economic Review*, 80(3), 313–336.
- Angrist, Joshua D., and William N. Evans, 1998, “Children and their Parents’ Labor Supply: Evidence from Exogenous Variation in Family Size”, *American Economic Review*, 88(3), 450–77.
- Angrist, Joshua D., and Alan B. Krueger, 1991, “Does Compulsory School Attendance Affect Schooling and Earnings”, *Quarterly Journal of Economics*, 106(4), 976–1014.
- Angrist, Joshua D., and Victor Lavy, 1999, “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement”, *Quarterly Journal of Economics*, 114(2), 533–75.
- Ashenfelter, Orley, 1987, “The Case for Evaluating Training Programs with Randomized Trials”, *Economics of Education Review*, 6(4), 333–338.
- Ashenfelter, Orley, and Daniel Hosken, 2010, “The Effect of Mergers on Consumer Prices: Evidence from Five Mergers on the Enforcement Margin”, *Journal of Law & Economics*, 53(3), 417–66.
- Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan, 2009, “The Miracle of Microfinance? Evidence from a Randomized Evaluation”, BREAD Working Paper 278.
- Barro, Robert J., and Charles J. Redlick, 2011, “Macroeconomic Effects from Government Purchases and Taxes”, *Quarterly Journal of Economics*, 126(1), 51–102.
- Berry, Steven, James Levinsohn and Ariel Pakes, 1995, “Automobile Prices in Market Equilibrium”, *Econometrica*, 63(4), 841–90.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, “How Much Should We Trust Differences-in-Differences Estimates?”, *Quarterly Journal of Economics*, 119(1), 249–75.
- Bleakley, Hoyt, 2007, “Disease and Development: Evidence from Hookworm Eradication in the American South”, *Quarterly Journal of Economics*, 122(1), 73–117.
- Bowers, William J., and Glenn L. Pierce, 1975, “The Illusion of Deterrence in Isaac Ehrlich’s Research on Capital Punishment”, *Yale Law Journal*, 85(2), 187–208.
15. Bronars, Stephen G., and Jeff Grogger, 1994, “The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment”, *American Economic Review*, 84(5), 1141–156.

- Bubb, Ryan, and Alex Kaufman, 2009, "Securitization and Moral Hazard: Evidence from a Lender Cutoff Rule", Federal Reserve Bank of Boston Public Policy Discussion Paper No. 09-5.
- Card, David, 1990, "The Impact of the Mariel Boatlift on the Miami Labor Market", *Industrial and Labor Relations Review*, 43(2), 245-57.
18. Card, David, 1995, "Earnings, Schooling and Ability Revisited," in *Research in Labor Economics*, Volume 14. Solomon W. Polachek, ed. Greenwich, Conn. and London, JAI Press, 23-48.
- Card, David, and Alan B. Krueger, 1992a, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the U. S. ", *Journal of Political Economy*, 100(1), 1-40.
- Card, David, and Alan B. Krueger, 1992b, "School Quality and Black-White Relative Earnings: A Direct Assessment", *Quarterly Journal of Economics*, 107(1), 151-200.
- Coleman, James S., et al, 1966, *Equality of Educational Opportunity*, Washington DC, U. S., Government Printing Office.
- Donohue, John J., and Justin Wolfers, 2005, "Uses and Abuses of Empirical Evidence in the Death Penalty Debate", *Stanford Law Review*, 58, 791-845.
23. Duflo, Esther, 2001, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", *American Economic Review*, 91(4), 795-813.
- Ehrlich, Isaac, 1975a, "The Deterrent Effect of Capital Punishment: A Question of Life and Death", *American Economic Review*, 65(3), 397-417.
- Ehrlich, Isaac, 1975b, "Deterrence: Evidence and Inference", *Yale Law Journal*, 85(2), 209-27.
- Ehrlich, Isaac, 1977a, "The Deterrent Effect of Capital Punishment: Reply", *American Economic Review*, 67(3), 452-58.
- Ehrlich, Isaac, 1977b, "Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence", *Journal of Political Economy*, 85(4), 741-88.
- Ehrlich, Isaac, and Zhiqiang Liu, 1999, "Sensitivity Analyses of the Deterrence Hypothesis: Let's Keep the Econ in Econometrics", *Journal of Law & Economics*, 42(1), 455-87.
- Erdem, Tulin, Susumu Imai, and Michael P. Keane, 2003, "Brand and Quantity Choice Dynamics under Price Uncertainty", *Quantitative Marketing and Economics*, 1(1), 5-64.
30. Evans, William N., and Jeanne Ringel, 1999, "Can Higher Cigarette Taxes Improve Birth Outcomes", *Journal of Public Economics*, 72(1), 135-54.
- Fehr, Ernst, and Lorenz Goette, 2007, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment", *American Economic Review*, 97(1), 298-317.
- Gertler, Paul, 2004, "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment", *American Economic Review*, 94(2), 336-41.
- Griliches, Zvi, 1986, "Economic Data Issues", In *Handbook of Econometrics*, vol. 3, ed. Zvi Griliches, and Michael D. Intriligator, 1465-1514, Amsterdam, North-Holland.
34. Gruber, Jonathan, 2000, "Disability Insurance Benefits and Labor Supply", *Journal of Political Economy*, 108(6), 1162-183.
- Gujarati, Damodar, 1978, *Basic Econometrics*, 1st ed. New York, McGraw-Hill.
- Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw, 2001, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica*, 69(1), 201-09.

- Hansen, Christian, 2007, "Asymptotic Properties of a Robust Variance Matrix Estimator for Panel Data when T is Large", *Journal of Econometrics*, 141(2), 597–620.
- Hastings, Justine S., 2004, "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California", *American Economic Review*, 94(1), 317–28.
39. Imbens, Guido W. and Wibert van der Klaauw, 1995, "Evaluating the Cost of Conscription in the Netherlands", *Journal of Business & Economic Statistics*, 13(2), 207–215.
- Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles, 2006, "Household Expenditure and the Income Tax Rebates of 2001", *American Economic Review*, 96(5), 1589–1610.
- Keys, Benjamin, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans", *Quarterly Journal of Economics*, 125(1), 307–62.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, 2007, "Experimental Analysis of Neighborhood Effects", *Econometrica*, 75(1), 83–119.
- Krueger, Alan B., 1999, "Experimental Estimates of Education Production Functions", *Quarterly Journal of Economics*, 114(2), 497–532.
- LaLonde, Robert J., 1986, "Evaluating the Econometric Evaluations of Training Programs with Experimental Data", *American Economic Review*, 76(4), 604–620.
- Leamer, Edward E., 1983, "Let's Take the Con Out of Econometrics", *American Economic Review*, 73(1), 31–43.
- Lee, David S., 2008, "Randomized Experiments from Non-random selection in U. S. House Elections", *Journal of Econometrics*, 142(2), 675–97.
47. Levitt, Steven D., 1997, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime", *American Economic Review*, 87(3), 270–90.
- Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz, 1987, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment", *American Economic Review*, 77(3), 251–77.
49. McClellan Mark, Barbara J. McNeil and Joseph P. Newhouse, 1994, "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?", *Journal of the American Medical Association*, 272(11), 859–66.
- Miguel, Edward, and Michael Kremer, 2004, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities", *Econometrica*, 72(1), 159–217.
- Moreira, Marcelo J., 2003, "A Conditional Likelihood Ratio Test for Structural Models", *Econometrica*, 71(4), 1027–48.
- Nevo, Aviv, 2000, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry", *The RAND Journal of Economics*, 31(3), 395–442.
- Passell, Peter, and John B. Taylor, 1977, "The Deterrent Effect of Capital Punishment: Another View", *American Economic Review*, 67(3), 445–51.
- Persson, Torsten, and Guido Tabellini, 2008, "The Growth Effect of Democracy: Is It Heterogeneous and How Can It Be Estimated?", Chap. 13 in *Institutions and Economic Performance*, ed. E. Helpman. Cambridge, MA, Harvard University Press.
- Powell, James L., 1994, "Estimation of Semiparametric Models", Chap. 41 in *Handbook of Econometrics*, vol. 4, ed. Robert Engle and Daniel McFadden, Amsterdam, North Holland.

- Ramey, Valerie, 2011, “Identifying Government Spending Shocks: It’s All in the Timing”, *Quarterly Journal of Economics*, 126(1), 1–50.
- Ramey, Valerie, and Matthew D. Shapiro, 1998, “Costly Capital Reallocation and the Effects of Government Spending”, *Carnegie Rochester Conference on Public Policy*, 48(1), 145–94.
- Richardson, Gary, and William Troost, 2009, “Monetary Intervention Mitigated Banking Panics during the Great Depression: Quasi-Experimental Evidence from a Federal Reserve District Border, 1929–1933”, *Journal of Political Economy*, 117(6), 1031–73.
- Rodrik, Dani, and Romain Wacziarg, 2005, “Do Democratic Transitions Produce Bad Outcomes?”, *American Economic Review*, 95(2), 50–55.
- Romer, Christina D., and David H. Romer, 1989, “Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz”, *NBER Macroeconomics Annual*, 4, 121–70.
- Romer, Christina D., and David H. Romer, 2004, “A New Measure of Monetary Shocks: Derivation and Implications”, *American Economic Review*, 94(4), 1055–1084.
- Romer, Christian D., and David H. Romer, 2010, “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks”, *American Economic Review*, 100(3), 763–801.
- Schultz, T. Paul, 2004, “School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program”, *Journal of Development Economics*, 74(1), 199–250.
- Sims, Christopher A., 1980, “Macroeconomics and Reality”, *Econometrica*, 48(1), 1–48.
- Sims, Christopher A. and Tao Zha, 2006, “Does Monetary Policy Generate Recessions?”, *Macroeconomic Dynamics*, 10(2), 231–72.
- Smets, Frank, and Raf Wouters, 2003, “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area”, *Journal of the European Economic Association*, 1(5), 1123–75.
67. van der Klaauw, Wilbert, 2002, “Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach”, *International Economic Review*, 43(4), 1249–1287.
- Velde, Francois, 2009, “Chronicles of a Deflation Unforetold”, *Journal of Political Economy*, 117(4), 591–634.
- White, Halbert, 1980, “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica*, 48(4), 817–38.